DOCUMENT RESUME

ED 181 070 TM 009 993

AUTHOR	Ebel, Robert L.
TITLE	The Paradox of Educational Testing.
INSTITUTION	National Council on Measurement in Education, East Lansing, Mich.
PUB DATE	76
NOTE	9p.
JOURNAL CIT	NCME Measurement in Education: v7 n4 Fall 1976
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Achievement Tests; Certification; Curriculum; Educational Accountability; Educational Quality; *Educational Testing; Minimum Competency Testing; Occupational Tests; *Public Opinion; Standardized Tests; Test Bias; *Testing Problems; Test Validity
IDENTIFIERS	Test Anxiety

ABSTRACT

There is currently a conflict between educational accountability and the distrust of standardized testing. Concern for the quality of education is based on evidence of students' academic dificiencies, the decline in test scores, and increasing education costs and school taxes. As a result of criticism, states have mandated, or are considering, competency testing programs. Occupational competency is also being certified more widely. Concurrently, criticism of testing programs is being voiced mainly by three groups: professional educators; reformers, and free-lance expose writers. Some of the most frequent criticisms include invalid tests; reliance on objective tests; labeling students; the influence of external testing programs; test anxiety; test bias; and norm-referenced test emphasis rather than use of criterion-referenced tests. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Vol. 7, No. 4

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY NCME TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED181070

NCME

measurement in education

A SERIES OF SPECIAL REPORTS OF THE NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

The Paradox of Educational Testing[1]

Robert L. Ebel
Professor of Education and Psychology
Michigan State University

ABOUT THIS REPORT

Robert L. Ebel

The furor over standardized testing continues. Relevant or irrelevant, invidious or enlightening, valid or invalid, biased or unbiased, the verbal barrages continue. In this issue of ME and in some of the upcoming issues as well we will present the thoughts and recommendations of some of the experts within the measurement community. The focus of these papers will be issues faced not only by the measurement specialist but by school personnel as well. ME is meant to serve the practitioner and, as such, we will continue to attempt to provide expert commentary in language suitable for "front-line" personnel. In this way we are hoping to alter the not-always-inaccurate perception of the university professor as a recluse in an ivy covered tower having little, if any, contact with the very people whose responsiblility it is to administer, interpret, and sometimes defend tests about which they know very little.

Dr. Robert Ebel has long been associated with excellence in the measurement profession. He has been affiliated with the State University of Iowa, Educational Testing Service, and is currently Professor of Education and Psychology at Michigan State University. He has served as President of American Educational Research Association, National Council on Measurement in Education, and Division 5 of the American Psychological Association.

PSR

The Paradox

We venture to predict that the mode of examination, by printed questions and written answers, will constitute a new era in the history of our schools.

Horace Mann

There is a paradox in educational measurement today. While assessments of achievement and competence are being more urgently called for and more widely employed than ever before, tests are, at the same time, being more sharply criticized and more strongly opposed. Perhaps these apparent inconsistencies are, in fact, wholly consistent. For, if the demands for more measurement of educational outcomes arises out of dissatisfaction with how much is being learned, those responsible for the teaching are likely to feel threatened by the possible public exposure of educational shortcomings. In any case, despite its expanding role in education measurement is under attack. Let us look briefly at contemporary demands for more evidence of achievement and competence, and then consider the arguments being made against the expansion, or indeed even the continued use of measurement in education.

Are They Learning Enough?

Criticism of schools and teaching is not new. It did not begin with Horace Mann, though he was one of the more prominent critics of American schools in the first half of the nineteenth century. He was also one of the first to make significant use of educational measurements to assess the effectiveness of schooling. Quite predictably, the written examinations he advocated were strongly opposed by the Boston schoolmasters whose pupils were to be examined.

[1] This material is taken from a forthcoming edition of Principles of Educational Measurement by Robert L. Ebel to be published by Prentice-Hall, Inc., in 1979.

TM009 993

In the decade of the 1970's, concern for the effectiveness of American schools became much stronger and more widespread than it had been before. Part of this concern arose from evidence, incidental and systematic, of deficiencies in learning. Some pupils given high school diplomas, were able to read only with difficulty or not at all. Incompetence in arithmetic, spelling, composition, geography, science, and other subjects of study were commonly reported. A sharp decline in the college admission test scores of high school students received wide publicity. It became apparent that in many schools the environment for learning had deterioriated badly. Discipline and vandalism became major problems. Hard study, with homework and rigorous grading of assignments was no longer fashionable.

Coupled strangely with declining pupil achievements were sharply increased school costs. Per pupil expenditures for instruction rose much faster than the general rate of inflation. Salaries for teachers in most states increased more rapidly than the average of all wages and salaries. No doubt some of the increase in teachers' salaries was a well warranted catch-up for a previously underpaid profession. No doubt the growing strength and militancy of teachers associations (unions) had much to do with the increase in salaries and the resultant increase in school costs. But the net effect of these increases was a correspondingly sharp increase in school taxes. It is not surprising that citizens, grumbling about taxes, began to ask for evidence that these tax dollars were doing some good. How much, they asked, are we getting for our money?

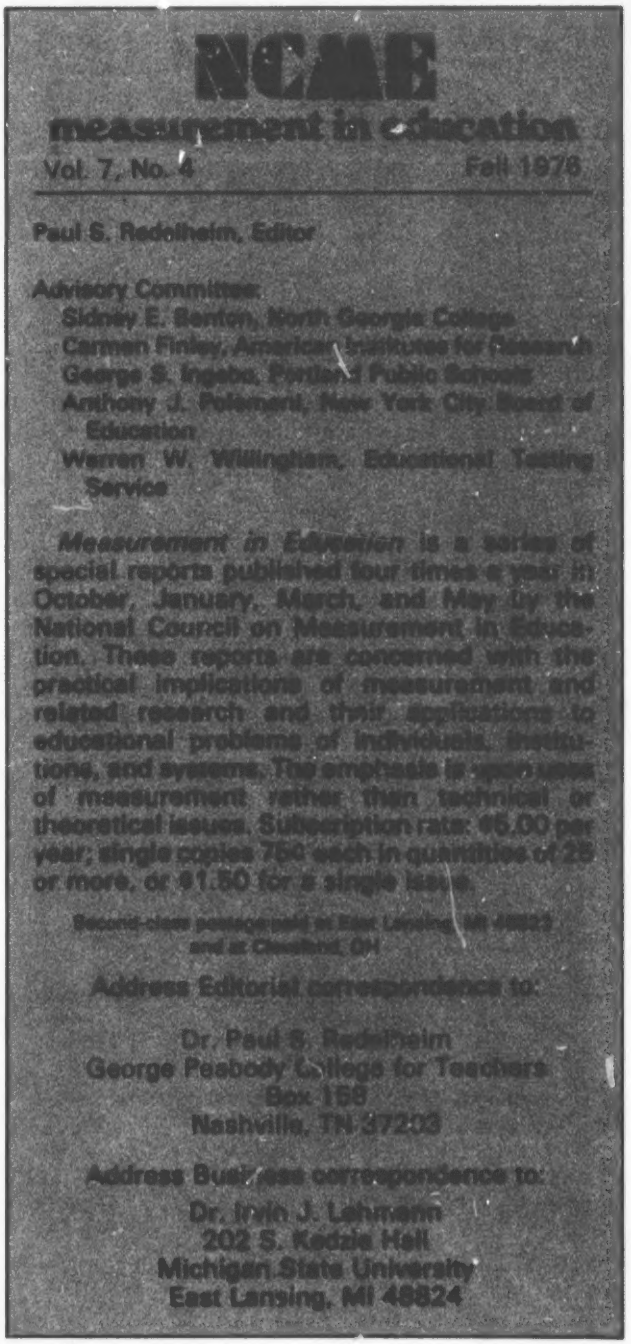
Mandated Assessment

In response to citizen concerns over the quality of student learning and the costs of providing instruction, a number of state legislatures passed laws that required tests to be given to assess the educational achievements of pupils in the public schools. Statewide educational testing began long ago. In the 1930's optional testing programs, often operated by state universities but sometimes by state departments of education, were available to schools in many of the states. But these programs were voluntary. Their stated purpose was to help schools to teach better, not to check up on how well they were doing. School administrators and school teachers are much more favorably inclined toward tests used to help them to do a better job than toward tests used to determine how well they have done it.

Some states have mandated programs of testing to determine the competencies of their high school graduates. The intent is to use this information to provide diplomas that will show specifically what the graduate has learned to know or to do as a result of schooling. This is a move in the direction of the British system of certifying A-level (advanced) or O-level (ordinary) passes on examinations in specific subject areas.

Accountability

The move toward mandated assessment is motivated by contemporary concerns for accountability. Hopeful parents and hard pressed tax payers believe that it is necessary and proper to hold school


NCME
measurement in education
Vol. 7, No. 4 Fall 1978

Paul S. Redelheim, Editor

Advisory Committee:
Sidney E. Benton, North Georgia College
Carmen Finley, American Institutes for Research
George S. Ingebo, Portland Public Schools
Anthony J. Polemeni, New York City Board of Education
Warren W. Willingham, Educational Testing Service

Measurement in Education is a series of special reports published four times a year in October, January, March, and May by the National Council on Measurement in Education. These reports are concerned with the practical implications of measurement and related research and their applications to educational problems of individuals, institutions, and systems. The emphasis is upon uses of measurement rather than technical or theoretical issues. Subscription rate: $6.00 per year; single copies 75¢ each in quantities of 25 or more, or $1.50 for a single issue.

Second-class postage paid at East Lansing, MI 48823 and at Cleveland, OH

Address Editorial correspondence to:

Dr. Paul S. Redelheim
George Peabody College for Teachers
Box 168
Nashville, TN 37203

Address Business correspondence to:

Dr. Irwin J. Lehmann
202 S. Kedzie Hall
Michigan State University
East Lansing, MI 48824

3

board members, school administrators, teachers, and pupils accountable for their successes or failures in the promotion of learning. School officials, on their part, tend to be wary of accountability. They fear, not without justification, that they may be held responsible for disappointing outcomes which were not their fault, for failures they could not prevent. However should this exempt them from accountability? Political leaders, business managers, football coaches, surgeons, and battlefield commanders are all exposed to similar hazards. They may hope to be judged fairly, but they can not escape judgment. Neither can educators. Surely the achievements of their pupils in learning is part of the evidence required to make a fair judgment.

Instructional Testing

Several recent developments in systematic instruction call for increased use of tests. Competency-based education, directed toward attainment of specific, behaviorally defined objectives of instruction, requires separate tests of the attainment of each of the competencies. Programmed instruction requires repeated testing to determine whether a pupil is ready to advance to the next phase of the program or needs to be "recycled" through the preceding phase. Indeed, in many programs for instruction, testing is not separate from but is imbedded in the instructional program. Individually guided education requires pretests to determine initial status and individually administered unit tests to determine progress along the guiding pathway.

Good teachers have always used tests to measure the achievements of individual pupils in learning and the effectiveness of instructional programs. Good students have always tested themselves for understanding and for recall as part of their methods of study. What has happened recently is that the testing has become a more prominent, a more systematically organized element of the instructional process. Such testing is obviously essential. There have been no calls for a moratorium on instructional testing, as there have been on standardized testing. But the same skills of test planning, question formulation, scoring, and score interpretation are basic to both.

Occupational Competency

Tests have long been used as part of the basis for certifying the competence of lawyers (bar exams), of doctors (medical boards), or accountants (CPA exams), of clinical psychologists, of nurses, of policemen and firemen, or real estate brokers, of barbers and hair-dressers, and of other professional practicioners and tradesmen. The list is constantly expanding. Dissatisfaction with the quality of some automobile repairs led to a movement recently to examine and certify auto mechanics.

City, state, and national civil service boards and commissions prepare, administer, and score enormous numbers of competitive examinations to help them select the best qualified government employees. In England an agency known as The City and Guilds of London Institute examines applicants for certification of competence in a wide range of occupational specialties. Possession of a City and Guilds certificate is a considerable asset, if not a prerequesite to employment in many fields.

What the tests for occupational competence can measure is knowledge of how a job ought to be done and, in a few cases, skill in doing it. This aspect of competence is surely important but just as surely it does not guarantee success in the practice of the profession or the trade. That will depend on the practicioners hard work, personality, good luck, and possibly other factors. In most occupations scores on tests of competence are related to success in practice. That the scores do not predict success infallibly may be due partly to shortcomings of the tests. But other elements of success, untested and probably untestable, account for much of the error in prediction.

As occupations become more specialized, as demands for competence grow in our complex society, tests of occupational competency will be used more and more. Good tests are urgently needed, and competent test specialists are urgently needed to supply them.

Criticisms Of Tests And Testing

No one who is well acquainted with the tests used in education is likely to claim that they are above criticism. For the process of assessing achievement in learning is much more complicated than it may seem to be at first glance. The naive view is that all one has to do is to give a test. Each student eithe passes it or fails it. The proportion who pass indicated the effectiveness of the educational program. Where the test comes from, what it actually measures, how good it is, what the scores mean, how the passing score is determined — these and many other hard questions that trouble test specialists a great deal trouble the general public hardly at all.

So very little of the criticism of tests and testing comes from the general public. Most of it comes from members of three special interest groups:

1. Professional educators who are uneasy about the accountability associated with standardized tests and external testing in general.

2. Reformers who regard testing as part of an

unsuccessful and outmoded instructional process.

3. Free-lance writers whose best-sellers purport to expose scandals in important human institutions.

From none of these groups of critics is one likely to get constructive criticisms of tests and testing. What they want is not to improve the tests, to correct their faults, to make them more useful educational tools. It is rather to discredit tests, to minimize their influence, or to get rid of them altogether.

None the less, the criticisms need to be taken seriously. For, if they are repeated often enough and loudly enough with no response, they will come to be accepted as true, no matter how little evidence there is to support them.

Here are some of the most frequently mentioned criticisms of tests and testing.

1. That the tests are *invalid,* measuring only superficial, unimportant aspects of achievement.

2. That *objective tests* are used mainly because they can be easily scored, regardless of serious shortcomings of triviality, ambiguity, and guessing.

3. That I.Q. *testing* has led to the labeling of particular puplis as bright or dull, in both cases distorting their expectations and diminishing their efforts.

4. That external testing programs impose *external controls* on the curriculum and lead teachers to give up sound long-range instruction in order to "teach to the test."

5. That testing places students under harmful *stress* and exposes them to unnecessary experiences of failure, destroying their self-confidence and killing the joy of learning.

6. That the tests carry a white, middle-class *bias,* misrepresenting the achievements and the potential of cultural minorities.

7. That conventional, single-score, norm-referenced, standardized tests should be discarded and be replaced by *criterion referenced tests* that indicate, not whether Johnny knows more than Mary, but exactly which objectives of instruction have been achieved and which have not.

Responses To The Criticisms

Regarding the matter of test validity it need only be said that tests provide the best information teachers and puplis can get on the success of efforts to teach and to learn.

Insofar as the values and limitations of objective tests are concerned with only a very few special exceptions, whatever can be tested at all can be tested with objective tests, usually with greater precision and often with greater convenience.

Some of the sharpest attacks on the tests used in education have been directed at intelligence, or IQ tests. The focus of recent controversy has been on whether or not there are racial differences in intelligence. This is part of the larger question of the extent to which a person's mental abilities, superior or deficient, are the results of heredity. Two observations may suffice for present purposes. The first is that IQ testing is probably more vulnerable to informed criticism than are most other forms of educational testing. The nature of intelligence is commonly misunderstood, and IQ's are commonly misinterpreted. They can be and have been used to discourage some pupils from trying to learn, and to excuse teachers from trying to teach them. The second observation is that there are important differences between intelligence tests and standardized tests of achievement in what is to be measured, how well it can be measured, and what effect its measurement will have on the progress of learning. Thus it is not true, as critics of testing sometimes imply, that the deficiences and abuses of the one are necessarily also characteristic of the other.

External Control Of Curricula

Standardized tests of achievement are sometimes attacked because of their supposed effect on the curricula of local schools. It is charged that they enforce curricular conformity, that they fail to test what local schools or particular teachers have been trying to teach, that they hamper curricular innovations or the use of open curricula. There is substance to these charges, but the effects are neither so overpowering nor so harmful as the critics imply.

If a school wishes to do well on certain tests it must of course see to it that the pupils receive instruction in the achievements sampled by the tests. But these are areas that panels of expert teachers have identified as important for most pupils to learn. A local school or a local teacher that decides to concentrate on achievements other than these should be prepared to do two things: (1) to argue convincingly to the public that what they are trying to teach is indeed more important for these pupils to

know than the things the standardized test is testing; and (2) to prepare, give, validate, and report the results on a local test that does measure what they have been trying to teach. Few of those who complain about the curricular restrictions of standardized testing are prepared to do either of these things.

A school that is teaching what the tests test will surely teach many other things beside. Even in the basic skill or core areas that the test does sample, there will be class time and teacher time to venture into interesting and useful areas of learning not covered by the standardized tests. Standardized tests can donimate local curricula only to the extent that school administrators and school teachers allow them to.

In some programs of open education there are covert goals for learning, and classroom activities are covertly manipulated so that pupils make systematic progress toward these goals. In such programs, standardized tests of achievement may be not only appropriate, but essential. Other programs have no particular goals for pupil learning, and are satisfied merely with maximum pupil freedom, trusting nature alone to do the job that others employ the art and science of teaching to help nature to do. That standardized tests of achievement are inappropriate for such programs may imply more criticism of the program than of the tests.

Teaching To The Test

If the effectiveness of instruction is to be assessed on the basis of student performance on a test, the temptation is strong for the instructor to prepare students to handle the specific questions that will be included in the test. This is often called "teaching to the test." Obviously, it is not an educationally beneficial procedure. Obviously, it tends to spoil the test as a measure of general achievement in the subject of the test. But if the test is a readily available, published test, an insecure or short-sighted teacher is quite likely to undertake to teach to the test. Even a test specially constructed for a particular assessment task is, in effect, a "published" test once it has been used. If teachers care enough, they can find out a great deal about the particular questions included in it. So it too may invite "teaching to the test."

Several things can be done to lessen the likelihood of teaching to the test. Teachers can be warned to avoid it and informed that their supervisors will be altert to notice it if it occurs. Pupils can be advised of its undesirability and asked to report its occurrance. Of course, the surest but most expensive way to forestall it is to prepart a new test for each new assessment effort.

Before leaving this subject, let us make an important distinction between "teaching to the test," i.e., attempting to fix in pupils minds the answers to particular test questions, and teaching material covered by the test, i.e., attempting to give pupils the capability of answering questions *like* those in the test on topics covered by the test. The first is thoroughly reprehensible. The second reflects purposeful teaching. Just as there is no warrant for giving away the answers to particular questions so there is no warrant for testing pupil performance on tasks they were never taught to perform. A teacher or school whose work is to be assessed is entitled to know what the pupils will be expected to do. This calls for close cooperation and clear communication between teachers on the one hand and test constructors on the other. An assessment test must be thoroughly relevant to the instruction it is intended to assess. Since a test can never elicit more than a sample of performance, much more will usually be taught than can be tested. But, the test should never go beyond what has been taught.

Does Testing Harm Pupils?

Several claims are made to support the charge that standardized testing harms pupils: that such tests threaten and upset pupils; that some pupils even break down and cry when faced with a test; that if a pupil gets a low score on a test he will become discouraged and quit trying; that his self concept will be seriously damaged; that standardized testing is incompatible with educational procedures designed to support the child.

There is, no doubt, anecdotal evidence to support some of these claims as affecting particular pupils. Common sense suggests, however, that the majority of pupils are not harmed by testing. There is no substantial survey data that would contradict common sense on this matter. Teachers seem much more often concerned with pupils who don't care enough how well or how poorly they do on such tests, than with the relatively rare exceptional instances of pupils who seem to care too much.

It is normal and biologically helpful to be somewhat anxious when facing any real test in life, regardless of one's age. But it is also a necessary part of growing up to learn to cope with the kind of tests life inevitably brings. Of the many challenges to a child's peace of mind, caused by such things as angry parents, playground bullies, bad dogs, shots from the doctor, and things that go bump in the night, standardized tests must surely be among the least fearsome for most children. Unwise parental pressure can, in some cases elevate anxiety to harmful levels. But usually the child who breaks down in tears at the prospect of a test has problems of security, adjustment, and maturity which testing

did not create, and which can not be solved by eliminating tests.

A pupil who consistently gets low scores on tests of what that pupil has tried hard to learn is indeed likely to become discouraged. If this does happen, the school can not claim to be offering a good educational program, and the teacher can not claim to be doing a good job of teaching. Most low scores on tests, however, go to pupils who, for one reason or another, have not tried very hard to learn. In the opinion of the teachers of such pupils it is the trying rather than the testing that is most in need of correction.

Pupils do indeed deserve support from their teachers. But what needs to be supported is the pupil's person and potential, not his mistakes or shortcomings. If teachers support ignorance or wrongdoing, they betray their profession. In the area of school achievement, what teachers need to support is progress toward the goals of learning. Standardized tests can help them to provide that support.

The Problem Of Test Bias

Standardized tests of educational achievement have also been attacked for their alleged bias against cultural minorities, and against pupils with poor reading skills. The reason for the attack, at least in part, is the fact that such pupils tend to make lower scores on the standardized tests. But surely lower scores alone do not signify bias. If they did, every spelling test would be biased against poor spellers, and every typing test against persons who never learned to type.

That a pupil who does poorly on a particular test in English might do better if the test were in Spanish, or if the questions were presented orally, does not mean that the original test is biased against them. It simply means that they do not have much of the precise achievement that particular test measures. Its linguistic context is part of the test. The particularity of what the test measures does not constitute bias.

The score of a pupil on an achievement test indicates how successfully the test questions were answered under the conditions of the test. The reasonable assumption is usually made that the pupil would be equally successful with other tasks requiring the same knowledge or ability. But if the pupil is frightened of being tested, the test score may not be an accurate indication of his knowledge or ability. Experimental studies have shown that the scores of some pupils can be raised somewhat by changing the conditions of testing. But the increase is usually rather small, too small to support the hypothesis that it is the conditions of testing which accounts for the lower test score of minority pupils.

What of the pupil who is offended by being asked to take a test at all. Surely, in this case, the test score

REPORTS AVAILABLE

Back issues at *Measurement in Education* are available for 1.50 each per single issue, and for 75¢ each in quantities of 25 or more for a single issue.

Vol. 1, No. 1	*Helping Teachers Use Tests* by Robert L. Thorndike
No. 2	*Interpreting Achievement P-ofiles — Uses and Warnings* by Eric F. Gardner
No. 3	*Mastery Learning and Mastery Testing* by Samuel T. Mayo
No. 4	*On Reporting Test Results to Community Groups* by Alden W. Badal & Edwin P. Larsen
Vol. 2, No. 1	*National Assessment Says* by Frank B. Womer
No. 2	*The PLAN System for Individualizing Education* by John C. Flanagan
No. 3	*Measurement Aspects of Performance Contracting* by Richard E. Schutz
No. 4	*The History of Grading Practices* by Louise Witmer Cureton
Vol. 3, No. 1	*Using Your Achievement Test Score Reports* by Edwin Gary Joselyn & Jack C. Merwin
No. 2	*An Item Analysis Service for Teachers* by Willard G. Warrington
No. 3	*On the Reliability of Ratings of Essay Examinations* by William E. Coffman
No. 4	*Criterion-Referenced Testing in the Classroom* by Peter W. Airasian and George F. Madaus
Vol. 4, No. 1	*Goals and Objectives in Planning and Evaluation: A Second Generation* by Victor W. Doherty and Walter E. Hathaway
No. 2	*Career Maturity* by John O. Crites
No. 3	*Assessing Educational Achievement in the Affective Domain* by Ralph W. Tyler
No. 4	*The National Test-Equating Study in Reading (The Anchor Test Study)* by Richard M. Jaeger
Vol. 5, No. 1	*The Tangled Web* by Fred F. Harcleroad
No. 2	*A Moratorium? What Kind?* by William E. Coffman
No. 3	*Evaluators, Educators, and the Publics: A Detente?* by William A. Mehrens
No. 4	*Shall We Get Rid of Grades?* by Robert L. Ebel
Vol. 6, No. 1	*On Evaluating a Project: Some Practical Suggestions* by John W. Wick
No. 2	*Measuring Reading Achievement: A Case for Criterion — Referenced Testing and Accountability* by S. Jay Samuels and Glenace E. Edwall
No. 3	*Security in a Citywide Testing Program* by Anthony J. Polemeni
No. 4	*Dear Mama: Why don't they love me anymore?* by Thomas J. Fitzgibbon
Vol. 7, No. 1	*The Discrepancy Evaluation Model (Part I)* by Andres Steinmetz
No. 2	*The Discrepancy Evaluation Model (Part II)* by Andres Steinmetz
No. 3	*What Graduate and Professional School Students Think About Admission Tests* by Leonard L. Baird

is not an accurate indication of knowledge. But is this the fault of the test or of the student? Should we say that the test is biased against the pupil? Or is it the pupil who is biased against the test?

The possibility of bias in intelligence tests is much greater than it is in achievement tests. Scores on a test of general intelligence are supposed to indicate a person's ability to learn many different kinds of things from his success in having learned a few other things. If the examinees all had nearly equal opportunity to learn what the intelligence test requires, the indications of general intelligence may be reasonably accurate. If not, the test may yield seriously biased indications of intelligence.

Criterion Referenced Tests

In recent years a new kind of test, referred to as a criterion referenced test has become popular. These tests have also been referred to as objectives referenced or domain referenced tests. Their purpose is to indicate which, and how many, of certain specific objectives of instruction a particular pupil has attained. Instead of a score based on the number of questions answered correctly, a criterion referenced tests yields a listing of objectives attained. The items on this list can, of course, be counted, and usually will be counted. Thus a kind of score can be obtained from a criterion referenced test.

A pupil's performance on such a test is presumably to be judged against an ideal criterion of perfect attainment of all the objectives, rather than by comparing the pupil's success against that of all other similar pupils. Thus criterion referenced tests typically provide no norms, or distributions of scores obtained from similar pupils. Of course comparisons can be made, and ususally will be made, between the numbers of objectives attained by different pupils. So a kind of norm is likely to evolve.

A view held by some critics of conventional tests, is that criterion referenced tests represent a significant and generally applicable improvement in the conception and methodology of achievement testing. These advocates have suggested that conventional norm referenced testing should be abandoned in favor of the new type. Conventional test statistics, they say, are not applicable to criterion referenced tests, and more appropriate new statistics must be developed.

Other test specialists see a more limited role for criterion referenced tests. Such tests can be used effectively, they say, in educational programs whose goals for learning are distinct, and relatively few in number, as in elementary arithmetic. Even here norms, they feel, will be useful in score interpretation. Criterion referenced tests focus on much more complete attainment of a much smaller number of objectives. This may be advantageous when one is dealing with mastery of basic skills. Whether it is equally desirable in other areas of learning may be open to question.

While there are major differences between criterion referenced and conventional norm referenced tests, there are also substantial similarities. Both have essentially the same job to do, that is to measure achievement in learning. Both are related to objectives of instruction. Elements of quality are essentially the same for both. The individual test questions used in the two are indistinguishable. Hence it seems unlikely that the development of criterion referenced tests will start any very radical or very sweeping revolution in the measurement of educational achievement.

Concluding Comments

The need for good tests of achievement in learning or of competence to practice a profession is not likely to diminish. It is likely to grow. The cirticism of tests is not likely to disappear, nor should it. The tests we currently use are not as good, on the whole, as they need to be. Informed criticism can help to make them better. But critics who urge that testing be limited or abandoned advise our schools and our society badly. Even the imperfect tests we now use serve us far better, in general, than no tests would do. Educators should work to improve, not to destroy testing.

Guide For School Testing Programs

Edited by:

Annie W. Ward
Margaret E. Backman
Bruce W. Hall
Joseph L. Mazur

Published by:

The National Council on Measurement in Education

From the Preface:

"We feel that the information in this monograph can be useful to a system-wide test director in organizing and operating a system-wide testing program. It should also be useful to superintendents, supervisors of instruction, principals, guidance directors, and teachers in helping them to understand what the test director is (or should be) doing and how the testing program relates to the rest of the school program."

The Chapters:

The Test Director and Testing Programs in Today's Schools

Organization of Testing Services

The Standardized Testing Program

Using Results of Standardized Tests

Teacher-made Tests: The Test Director's Role

Reporting Test Results to the Public

This 88-page softbound monograph is available from:

"A GUIDE TO TESTING PROGRAMS"
Order Form

Name _____ Address _____

Address _____ _____

	No.	Cost
1-5 Copies @ $1.75 each*	_____	_____
6-49 Copies @ $1.50 each	_____	_____
50+ Copies @ $1.25 each	_____	_____

NOTE: Payment in U.S. funds only.

Nat'l. Council on Measurement in Education
202 S. Kedzie Hall, Mich. State Univ.
E. Lansing, MI 48824

*Single non-member, library and institutional orders are $3.50 per copy.
Invoiced orders ADD $1.00 per order.

NCME NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION

202 S. Kedzie Hall
Michigan State University
East Lansing, Michigan 48824

Second class postage paid
at East Lansing, Mich.
48823 and
additional mailing points.